



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : C12N 15/12, C07K 14/47, C12Q 1/68, A61K 38/17, G01N 33/50, 33/53		A1	(11) International Publication Number: WO 00/66728
			(43) International Publication Date: 9 November 2000 (09.11.00)
(21) International Application Number: PCT/IL00/00252 (22) International Filing Date: 3 May 2000 (03.05.00) (30) Priority Data: 129734 3 May 1999 (03.05.99) IL (71) Applicant (for all designated States except US): COMPUGEN LTD. [IL/IL]; 72 Pinchas Rosen Street, 69512 Tel Aviv (IL). (72) Inventors; and (75) Inventors/Applicants (for US only): SAVITZKY, Kinneret [IL/IL]; 44 Metodela Street, 69548 Tel Aviv (IL). MINTZ, Liat [IL/IL]; 14 Harav Kook Street, 47254 Ramat Hasharon (IL). (74) Agent: REINHOLD COHN AND PARTNERS; P.O. Box 4060, 61040 Tel-Aviv (IL).		(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	
(54) Title: STAR HOMOLOGUES			
(57) Abstract The invention concerns novel nucleic acid and amino acid sequences which are homologues to steroidogenic acute regulatory protein. The invention further concerns expression vectors comprising these sequences and host cells transfected with the vectors. The invention further concerns pharmaceutical compositions.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

STAR HOMOLOGUES

FIELD OF THE INVENTION

5 The present invention concerns novel nucleic acid sequences, vectors and host cells containing them, amino acid sequences encoded by said sequences, and antibodies reactive with said amino acid sequences, as well as pharmaceutical compositions comprising any of the above. The present invention further concerns methods for screening for candidate activator or deactivators utilizing said amino
10 acid sequences.

BACKGROUND OF THE INVENTION

The following publications are believed to be relevant as Background of the Invention:

- 15 (1) Douglas, M.S. and Clark, B.J., *Steroids*, 62:29-36 (1997).
- (2) Miller, W.L., *J. Steroid Biochem. Molec. Biol.*, 55(5,6):607-616 (1995).
- (3) Kallan, *et al.*, *Molecular and Cellular Endocrinology*, 145:39-45 (1998).
- 20 (4) Gharani, *et al.*, *Human Molecular Genetics*, 6(3):397-402 (1997).
- (5) Legro, R.S., *Molecular and Cellular Endocrinology*, 145:103-110 (1998).

Steroid hormones produced in the placenta, adrenals, ovaries, and gonads are important mediators of tissue differentiation, development, and homeostasis.
25 The synthesis of these important mediators of both genomic and non-genomic effects is tightly regulated in a temporal and tissue-specific manner. The acute

- 2 -

stimulation of steroidogenesis in the adrenals and gonads is triggered by trophic hormone-induced generation of cAMP with subsequent activation of phosphorylation and gene transcription by the cAMP-dependent protein kinase A (PKA). The rate-limiting step in steroid hormone synthesis is the side-chain
5 cleavage of cholesterol to form pregnenolone, a step catalyzed by the side-chain cleavage enzyme system (P450scc), which is located on the matrix side of the inner mitochondrial membrane.

The translocation of cholesterol from the cholesterol-rich outer mitochondrial membrane to the cholesterol-poor inner membrane is the true
10 rate-limiting step in the acute control of steroidogenesis. Consistent with this proposition is the finding that freely diffusible cholesterol analogs (hydroxysterols) that are able to reach the inner mitochondrial membrane promote maximal levels of steroidogenesis in the absence of trophic stimulation. It has also been shown that the acute steroidogenic response is blocked within minutes of treatment with
15 inhibitors of protein synthesis (i.e. cyclo-heximide or puromycin). Collectively, these observations suggested a model in which trophic hormone, through the intermediacy of cAMP, promotes the *de novo* synthesis of a labile or short-acting protein which functions to enhance transport of cholesterol from the outer to the inner mitochondrial membrane where it serves as substrate for the P450scc enzyme
20 system.⁽³⁾

The protein responsible for this transport is termed "*steroidogenic acute regulatory protein*" (StAR).

The importance of StAR in the regulation of steroidogenesis has been dramatically demonstrated in recent studies on the disease congenital lipoid adrenal
25 hyperplasia (lipoid CAH). Lipoid CAH is a lethal condition that results from a complete inability of the newborn infant to synthesize steroids. The lack of mineralocorticoids and glucocorticoids results in death within days to weeks of birth if not detected and treated with adequate steroid hormone and salt

- 3 -

replacement therapy. It was demonstrated that lipoid CAH is due to defects in StAR expression or function.⁽²⁾

The human StAR cDNA encodes a 285 amino acid protein with 37 kDa mobility on SDS polyacrylamide gels (SDS-PAGE). The StAR protein sequence is highly conserved (80-88% identity and 90% similarity) across all of the species thus far compared, including: the mouse, human, hamster, rat, cow, sheep and pig. The protein contains putative phosphorylation sites for several protein kinases including the protein kinase A (PKA), calmodulin- dependent kinase II (CAM kinase II), and protein kinase C (PKC). The amino terminal 33 amino acids of the predicted polypeptide possesses a high net positive charge and the potential to form an amphipathic helical structure typical of mitochondrial signal sequences. Both murine and human StAR proteins are efficiently imported and processed by mitochondria of steroidogenic and non-steroidogenic cells.⁽³⁾

Polycystic ovary syndrome (PCOS) is a highly prevalent endocrine disorder which is characterized by hyperandrogenaemia and represents the most common cause of anovulatory infertility and hirsutism. PCOS has been estimated to have a population prevalence of between 5-10%. The characteristic polycystic ovarian morphology, however, may be found in up to 22% of the normal population, with > 90% of these women having at least one mild symptom that may be considered a clinical marker of PCOS.

Hyperandrogenaemia is seen both in women with PCOS and men with premature male pattern baldness suggesting an underlying disorder of androgen biosynthesis or metabolism. Androgens are synthesized by the adrenals, the theca cell layer of the developing ovarian follicle and the testicular Leydig cells. Both scalp hair loss and hirsutism are known to be mediated by androgens. The sensitivity of the hair follicle to androgens is dependent on a number of factors, such as serum concentrations of bioavailable androgens and the presence and number of androgen receptors.⁽⁴⁾

- 4 -

It has been shown that theca cells from polycystic ovaries show a significant increase in both androstenedione and progesterone production *in vitro* when compared to normal theca. There have been indications that a marker, CYP11a, which has been localized to chromosome 15q23- was associated with the
5 Stien-Levental Syndrome (polycyclic ovary syndrome PCOS)⁽⁴⁾.

In addition, androgens have been known to play a major role in the regulation of various aspects of the biology of prostate cancer cells.

GLOSSARY

10 In the following description and claims use will be made, at times, with a variety of terms, and the meaning of such terms as they should be construed in accordance with the invention is as follows:

*“Steroidogenic acute regulatory protein-homolog (StAR-B) nucleic acid
15 sequence”* – the sequence shown in SEQ ID NO: 1, sequences having at least 70% identity to said sequence and fragments of the above sequences being 20 b.p. long. This sequence is a sequence coding for a novel homolog of the known StAR protein. The two sequences share homology at the C-terminus which is the region responsible for physiological and catalytic activity of StAR. However, the term
20 StAR-B does not necessarily signify that StAR-B protein coded by the above sequence has the same or even similar physiological effects as known StAR, merely that it shows sequence homology with the known StAR .

*“Steroidogenic acute protein (StAR-B product) – also referred at times as the
25 “StAR-B protein” or “StAR-B polypeptide”* – is an amino acid shown in SEQ ID NO: 2. The amino acid sequence may be a peptide, a protein, as well as peptides or proteins having *chemically modified* amino acids (see below) such as a glycopeptide or glycoprotein. An example of an StAR-B product is shown in SEQ ID NO: 2. The term also includes analogues of said sequences in which one or
30 more amino acids has been added, deleted, *substituted* (see below) or *chemically*

- 5 -

modified (see below) as well as fragments of this sequence having at least 10 amino acids.

5 “*Nucleic acid sequence*” – a sequence composed of DNA nucleotides, RNA nucleotides or a combination of both types and may includes natural nucleotides, chemically modified nucleotides and synthetic nucleotides.

10 “*Amino acid sequence*” – a sequence composed of any one of the 20 naturally appearing amino acids, amino acids which have been *chemically modified* (see below), or composed of synthetic amino acids.

“*Fragment of StAR-B product*” - a polypeptide which has an amino acid sequence which is the same as part of but not all of the amino acid sequence of the StAR-B product.

15 “*Fragments of StAR-B nucleic acid sequence*” a continuous portion, preferably of about 20 nucleic acid sequences of the StAR-B nucleic acid sequence.

20 “*Conservative substitution*” - refers to the substitution of an amino acid in one class by an amino acid of the same class, where a class is defined by common physicochemical amino acid side chain properties and high substitution frequencies in homologous proteins found in nature, as determined, for example, by a standard Dayhoff frequency exchange matrix or BLOSUM matrix. [Six general classes of amino acid side chains have been categorized and include: Class I (Cys); Class II (Ser, Thr, Pro, Ala, Gly); Class III (Asn, Asp, Gln, Glu);
25 Class IV (His, Arg, Lys); Class V (Ile, Leu, Val, Met); and Class VI (Phe, Tyr, Trp). For example, substitution of an Asp for another class III residue such as Asn, Gln, or Glu, is a conservative substitution.

- 6 -

"Non-conservative substitution" - refers to the substitution of an amino acid in one class with an amino acid from another class; for example, substitution of an Ala. a class II residue, with a class III residue such as Asp, Asn, Glu, or Gln.

5 *"Chemically modified"* - when referring to the product of the invention, means a product (protein) where at least one of its amino acid residues is modified either by natural processes, such as processing or other post-translational modifications, or by chemical modification techniques which are well known in the art. Among the numerous known modifications typical, but not exclusive examples include:
10 acetylation, acylation, amidation, ADP-ribosylation, glycosylation, GPI anchor formation, covalent attachment of a lipid or lipid derivative, methylation, myristylation, pegylation, prenylation, phosphorylation, ubiquitination, or any similar process.

15 *"Biologically active"* - refers to the StAR-B product which have, regulatory or biochemical functions on the same target sites which naturally occurring StAR-B influence, for example an effect in regulation of steroid synthesis, in particular androgen synthesis. This term may refer to the same biological activity as manifested in different cells in different tissues which are responsible for
20 hyperandrogenism, such as Theca cells, Leydig cells and cells of the adrenal cortex.

"Immunologically active" defines the capability of a natural, recombinant or synthetic StAR-B product, or any fragment thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.
25 Thus, for example, a biologically active fragment of StAR-B product denotes a fragment which retains some or all of the immunological properties of the StAR-B product, e.g. can bind specific anti-StAR-B product antibodies or which

- 7 -

can elicit an immune response which will generate such antibodies or cause proliferation of specific immune cells which produce StAR-B.

"Optimal alignment" - is defined as an alignment giving the highest percent identity score. Such alignment can be performed using a variety of commercially available sequence analysis programs, such as the local alignment program LALIGN using a ktup of 1, default parameters and the default PAM. A preferred alignment is the one performed using the CLUSTAL-W program from MacVector (TM), operated with an open gap penalty of 10.0, an extended gap penalty of 0.1, and a BLOSUM similarity matrix. If a gap needs to be inserted into a first sequence to optimally align it with a second sequence, the percent identity is calculated using only the residues that are paired with a corresponding amino acid residue (i.e., the calculation does not consider residues in the second sequences that are in the "gap" of the first sequence).

"Having at least X% identity" - with respect to two amino acid or nucleic acid sequence sequences, refers to the percentage of residues that are identical in the two sequences when the sequences are optimally aligned. Thus, 70% amino acid sequence identity means that 70% of the amino acids in two or more optimally aligned polypeptide sequences are identical.

"Isolated nucleic acid molecule having an StAR-B nucleic acid sequence" - is a nucleic acid molecule that includes the coding StAR-B nucleic acid sequence. Said isolated nucleic acid molecule may include the StAR-B nucleic acid sequence as an independent insert; may include the StAR-B nucleic acid sequence fused to an additional coding sequences, encoding together a fusion protein in which the StAR-B coding sequence is the dominant coding sequence (for example, the additional coding sequence may code for a signal peptide); the

- 8 -

StAR-B nucleic acid sequence may be in combination with non-coding sequences, e.g., introns or control elements, such as promoter and terminator elements or 5' and/or 3' untranslated regions, effective for expression of the coding sequence in a suitable host; or may be a vector in which the StAR-B protein coding sequence is a heterologous.

"Expression vector" - refers to vectors that have the ability to incorporate and express heterologous DNA fragments in a foreign cell. Many prokaryotic and eukaryotic expression vectors are known and/or commercially available. Selection of appropriate expression vectors is within the knowledge of those having skill in the art.

"Deletion" - is a change in either nucleotide or amino acid sequence in which one or more nucleotides or amino acid residues, respectively, are absent.

"Insertion" or "addition" - is that change in a nucleotide or amino acid sequence which has resulted in the addition of one or more nucleotides or amino acid residues, respectively, as compared to the naturally occurring sequence.

"Substitution" - replacement of one or more nucleotides or amino acids by different nucleotides or amino acids, respectively. As regards amino acid sequences the substitution may be conservative or non-conservative.

"Antibody" - refers to IgG, IgM, IgD, IgA, and IgG antibody. The definition includes polyclonal antibodies or monoclonal antibodies. This term refers to whole antibodies or fragments of the antibodies comprising the antigen-binding domain of the anti-StAR-B product antibodies, e.g. antibodies without the Fc

- 9 -

portion. single chain antibodies, fragments consisting of essentially only the variable, antigen-binding domain of the antibody, etc.

"Activator" - as used herein, refers to a molecule which mimics the effect of the natural StAR-B product or at times even increases or prolongs the duration of the biological activity of said product, as compared to that induced by the natural product. The mechanism may be by binding to the inner mitochondria membrane thus increasing the activity of StAR-B, by prolonging the lifetime of the StAR-B, by increasing the activity of the StAR-B on its target (transport of cholesterol), by increasing the affinity of StAR-B to moieties which it binds (such as cholesterol) etc. Activators may be polypeptides, nucleic acids, carbohydrates, lipids, or derivatives thereof, or any other molecules which can bind to and activate the StAR-B product.

"Deactivator" or (*"Inhibitor"*) - refers to a molecule which modulates the activity of the StAR-B product in an opposite manner to that of the activator, by decreasing or shortening the duration of the biological activity of the StAR-B product. This may be done by blocking the binding of the StAR-B to cholesterol (competitive or non-competitive inhibition), by causing rapid degradation of the StAR-B, etc. by inhibiting association of the StAR-B with the inner membrane of the mitochondria, etc. Deactivators may be polypeptides, nucleic acids, carbohydrates, lipids, or derivatives thereof, or any other molecules which bind to and modulate the activity of said product.

"Treating a disease" - refers to administering a therapeutic substance effective to ameliorate symptoms associated with a disease, to lessen the severity or cure the disease, or to prevent the disease from occurring.

- 10 -

“*Detection*” – refers to a method of detection of a disease. This term may refer to detection of a predisposition to a disease.

“*Probe*” – the StAR-B nucleic acid sequence, or a sequence complementary therewith, when used to detect presence of other similar sequences in a sample. The detection is carried out by identification of hybridization complexes between the probe and the assayed sequence. The probe may be attached to a solid support or to a detectable label.

SUMMARY OF THE INVENTION

10 The present invention is based on the surprising finding that there exist in humans a novel homolog of the StAR having a significant homology to the C-terminus of native StAR, which is the physiologically active region of that protein. The novel homolog was termed “*StAR-B*”. StAR-B was found to be localized on chromosome 15q23-q24. On this locus are also localized cytochrome 15 P450_{scc} variants, and the Stein Levental Syndrome (PCOS) was linked to this site⁽⁴⁾. It is now believed that modulation of StAR-B levels and/or activity underlies the pathologies associated with PCOS as well as other physiological conditions linked to PCOS families such as male premature baldness (MPB).

Thus the present invention provides by its first aspect, a novel isolated 20 nucleic acid molecule comprising or consisting of the sequence SEQ ID NO: 1, fragments of said sequence having at least 20 nucleic acids, or a molecule comprising a sequence having at least 70%, preferably 80%, and most preferably 90% identity to SEQ ID NO:1. The present invention further provides a protein or polypeptide comprising or consisting of an amino acid sequence encoded by any of 25 the above nucleic acid sequences, termed herein “*StAR-B product*”, for example, an amino acid sequence having the sequence as depicted in SEQ ID NO: 2, fragments of the above amino acid sequence having a length of at least 10 amino

- 11 -

acids, as well as homologs of the amino acid sequences SEQ ID NO.:2 in which one or more of the amino acid residues has been substituted (by conservative or non-conservative substitution) added, deleted, or chemically modified.

The present invention further provides nucleic acid molecule comprising or
5 consisting of a sequence which encodes the above amino acid sequences, (including the fragments and analogs of the amino acid sequences). Due to the degenerative nature of the genetic code, a plurality of alternative nucleic acid sequences, beyond SEQ ID NO:1, can code for the amino acid sequence of the invention. Those alternative nucleic acid sequences which code for the same amino
10 acid sequences codes by the sequence SEQ ID NO: 1 are also an aspect of the of the present invention.

The present invention further provides expression vectors and cloning vectors comprising any of the above nucleic acid sequences, as well as host cells transfected by said vectors.

15 The present invention still further provides pharmaceutical compositions comprising, as an active ingredient, said nucleic acid molecules, said expression vectors, or said protein or polypeptide.

These pharmaceutical compositions are suitable for the treatment of diseases and pathological conditions, which can be ameliorated or cured by raising the level
20 of the StAR-B product. Typically these are diseases which are manifested by non-normal levels of steroid hormones (which can be higher or lower than normal levels). The compositions are intended to restore the levels to normal levels. Thus the pharmaceutical compositions may serve as alternative regions for hormonal administration. Especially for the treatment of PCOS-involved conditions as well
25 as MPB By a second aspect, the present invention provides a nucleic acid molecule comprising or consisting of a non-coding sequence which is complementary to that of SEQ ID NO:1, or complementary to a sequence having at least 70% identity to said sequence or a fragment of said two sequences. The complementary sequence

- 12 -

may be a DNA sequence which hybridizes with the SEQ of ID NO:1 or hybridizes to a portion of that sequence having a length sufficient to inhibit the transcription of the complementary sequence. The complementary sequence may be a DNA sequence which can be transcribed into an mRNA being an antisense to the mRNA transcribed from SEQ ID NO:1 or into an mRNA which is an antisense to a fragment of the mRNA transcribed from SEQ ID NO:1 which has a length sufficient to hybridize with the mRNA transcribed from SEQ ID NO:1, so as to inhibit its translation. The complementary sequence may also be the mRNA or the fragment of the mRNA itself.

10 The nucleic acids of the second aspect of the invention may be used for therapeutic or diagnostic applications for example for detection of the expression of StAR-B in various tissues such as to ovary, adrenal, placenta, etc.

The present invention also provides expression vectors comprising any one of the above defined complementary nucleic acid sequences and host cells transfecting with said nucleic acid sequences or vectors, being complementary to those specified in the first aspect of the invention.

15 The invention also provides anti-StAR-B product antibodies, namely antibodies directed against the StAR-B product which specifically bind to said StAR-B product. Said antibodies are useful both for diagnostic and therapeutic purposes. For example said antibody may be as an active ingredient in a pharmaceutical composition as will be explained below.

20 The present invention also provides pharmaceutical compositions comprising, as an active ingredient, the nucleic acid molecules which comprise or consist of said complementary sequences, or of a vector comprising said complementary sequences. The pharmaceutical composition thus provides pharmaceutical compositions comprising, as an active ingredient, said anti-StAR-B product antibodies.

The pharmaceutical compositions comprising said anti-StAR-B product antibodies or the nucleic acid molecule comprising said complementary sequence,

- 13 -

are suitable for the treatment of diseases and pathological conditions where a therapeutically beneficial effect may be achieved by neutralizing the StAR-B or decreasing the amount of the StAR-B product or blocking its binding to its effector (cholesterol), for example, by the neutralizing effect of the antibodies, or by the decrease of the effect of the antisense mRNA in decreasing expression level of the StAR-B product. Mostly these diseases are manifested by non-normal level of steroid hormones in the diseased persons which may be regulated to produce normal levels by utilizing the pharmaceutical compositions of the invention. Thus the compositions of the invention may serve as alternative treatment regimes for hormonal administration. Preferably the pharmaceutical compositions are for the treatment of conditions involving PCOS and MPB. By another application, the therapeutics may be utilized for the treatment of prostate cancer.

According to the third aspect of the invention the present invention provides methods for detecting the level of the transcript (mRNA) of said StAR-B product in a body fluid sample, or in a specific tissue sample (notably the ovary, placenta, adrenal), for example by use of probes comprising or consisting of said coding sequences; as well as methods for detecting levels of expression of said product in tissue, e.g. by the use of antibodies capable of specifically reacting with the above amino acid sequences. Detection of the level of the expression of the StAR-B variant of the invention may be indicative of a plurality of physiological or pathological conditions.

The method, according to this latter aspect, for detection of a nucleic acid sequence which encodes the StAR-B product in a biological sample, comprises the steps of:

- (a) providing a probe comprising at least one of the nucleic acid sequence defined above;
- (b) contacting the biological sample with said probe under conditions allowing hybridization of nucleic acid sequences thereby enabling formation of hybridization complexes;

(c) detecting hybridization complexes, wherein the presence of the complex indicates the presence of nucleic acid sequence encoding the StAR-B product in the biological sample.

By a preferred embodiment the probe is part of a nucleic acid chip used for
5 detection purposes, i.e. the probe is a part of an array of probes each present in a known location on a solid support.

The nucleic acid sequence used in the above method may be a DNA sequence an RNA sequence, etc; it may be a coding or a sequence or a sequence complementary thereto (for respective detection of RNA transcripts or
10 coding-DNA sequences). By quantization of the level of hybridization complexes and calibrating the quantified results it is possible also to detect the level of the transcript in the sample.

Methods for detecting mutations in the region coding for the StAR-B product are also provided, which may be methods carried-out in a binary fashion,
15 namely merely detecting whether there is any mismatches between the normal StAR-B nucleic acid sequence and the one present in the sample, or carried-out by specifically detecting the nature and location of the mutation.

The present invention also concerns a method for detecting StAR-B product both for determining its presence ,as well as its level or alterations in its level in a
20 biological sample, comprising the steps of:

(a) contacting with said biological sample the antibody of the invention, thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

wherein the presence of said antibody-antigen complex correlates with the
25 presence of StAR-B product in said biological sample.

Both detection of StAR-B product and transcript, for example in urine samples, may be used to discriminate between various stages of prostate cancer. Changes in levels as compared to normal controls may be indicative of a pathological state.

- 15 -

By yet another aspect the invention also provides a method for identifying candidate compounds capable of binding to the StAR-B product and modulating its activity (being either activators or deactivators). The method includes:

(i) providing a protein or polypeptide comprising an amino acid sequence substantially as depicted in SEQ ID NO: 2, or a fragment of such a sequence:

(ii) contacting a candidate compound with said amino acid sequence;

(iii) measuring the physiological effect of said candidate compound on the activity of the amino acid sequences and selecting those compounds which show a significant effect on said physiological activity.

The activity of the amino acid which should be changed by the modulator (being either the activator or deactivator) may be for example the binding of the StAR-B product to cholesterol, effect on the transport rate of cholesterol to the inner mitochondrial membrane, effect on steroid synthesis, etc. Any modulator which changes such an activity has an intersecting potential, as serving as an actuator or deactivator.

The present invention also concerns compounds identified by the above methods described above, which compound may either be an activator of the StAR-B product or a deactivator thereof.

20

BRIEF DESCRIPTION OF THE DRAWINGS

In order to understand the invention and to see how it may be carried out in practice, a preferred embodiment will now be described, by way of non-limiting example only, with reference to the accompanying drawings, in which:

Fig. 1 is the nucleic acid sequence of the invention SEQ ID NO: 1;

Fig. 2 is the amino acid sequence of the invention SEQ ID NO: 2; and

Fig. 3 is a comparison of the known StAR sequence and the StAR-B sequence of the invention.

25

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

Example I: STAR-B - nucleic acid sequence

The nucleic acid sequences of the invention include nucleic acid
5 sequences which encode StAR-B product and fragments and analogs thereof. The
nucleic acid sequences may alternatively be sequences complementary to the
above coding sequence, or to a region of said coding sequence. The length of the
complementary sequence is sufficient to avoid the expression of the coding
sequence. The nucleic acid sequences may be in the form of RNA or in the form
10 of DNA, and include messenger RNA, synthetic RNA and DNA, cDNA, and
genomic DNA. The DNA may be double-stranded or single-stranded, and if
single-stranded may be the coding strand or the non-coding (anti-sense,
complementary) strand. The nucleic acid sequences may also both include
dNTPs, rNTPs as well as non naturally occurring sequences. The sequence may
15 also be a part of a hybrid between an amino acid sequence and a nucleic acid
sequence.

In a general embodiment, the nucleic acid sequence has at least 70%,
preferably 80% or 90% sequence identity with the sequence identified as SEQ ID
NO:1.

20 The nucleic acid sequences may include the coding sequence by itself. By
another alternative the coding region may be in combination with additional
coding sequences, such as those coding for fusion protein or signal peptides, in
combination with non-coding sequences, such as introns and control elements,
promoter and terminator elements or 5' and/or 3' untranslated regions, effective
25 for expression of the coding sequence in a suitable host, and/or in a vector or host
environment in which the StAR-B nucleic acid sequence is introduced as a
heterologous sequence.

- 17 -

The nucleic acid sequences of the present invention may also have the product coding sequence fused in-frame to a marker sequence which allows for purification of the StAR-B product. The marker sequence may be, for example, a hexahistidine tag to provide for purification of the mature polypeptide fused to the marker in the case of a bacterial host, or, the marker sequence may be a hemagglutinin (HA) tag when a mammalian host, e.g. COS-7 cells, is used. The HA tag corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson, I., *et al. Cell* 37:767 (1984)).

Also included in the scope of the invention are fragments also referred to herein as oligonucleotides, typically having at least 20 bases, preferably 20-30 bases corresponding to a region of the coding-sequence nucleic acid sequence. The fragments may be used as probes, primers, and when complementary also as antisense agents, and the like, according to known methods.

As indicated above, the nucleic acid sequence may be substantially depicted in SEQ ID NO:1 or fragments thereof or sequences having at least 70%, preferably 70-80%, most preferably 90% identity to the above sequence. Alternatively, due to the degenerative nature of the genetic code, the sequence may be a sequence coding the amino acid sequence of SEQ ID NO:2, or fragments or analogs of said amino acid sequence.

20

A. Preparation of nucleic acid sequences The nucleic acid sequences may be obtained by screening cDNA libraries using oligonucleotide probes which can hybridize to or PCR-amplify nucleic acid sequences which encode the StAR-B products disclosed above. cDNA libraries prepared from a variety of tissues are commercially available and procedures for screening and isolating cDNA clones are well-known to those of skill in the art. Such techniques are described in, for example, Sambrook *et al.* (1989) *Molecular Cloning: A Laboratory Manual* (2nd Edition), Cold Spring Harbor Press, Plainview, N.Y. and Ausubel FM *et al.*

25

(1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York, N.Y.

The nucleic acid sequences may be extended to obtain upstream and downstream sequences such as promoters, regulatory elements, and 5' and 3' untranslated regions (UTRs). Extension of the available transcript sequence may be performed by numerous methods known to those of skill in the art, such as PCR or primer extension (Sambrook *et al.*, *supra*), or by the RACE method using, for example, the Marathon RACE kit (Clontech, Cat. # K1802-1).

Alternatively, the technique of "restriction-site" PCR (Gobinda *et al.* *PCR Methods Applic.* 2:318-22, (1993)), which uses universal primers to retrieve flanking sequence adjacent a known locus, may be employed. First, genomic DNA is amplified in the presence of primer to a linker sequence and a primer specific to the known region. The amplified sequences are subjected to a second round of PCR with the same linker primer and another specific primer internal to the first one. Products of each round of PCR are transcribed with an appropriate RNA polymerase and sequenced using reverse transcriptase.

Inverse PCR can be used to amplify or extend sequences using divergent primers based on a known region (Triglia, T. *et al.*, *Nucleic Acids Res.* 16:8186, (1988)). The primers may be designed using OLIGO(R) 4.06 Primer Analysis Software (1992; National Biosciences Inc, Plymouth, Minn.), or another appropriate program, to be 22-30 nucleotides in length, to have a GC content of 50% or more, and to anneal to the target sequence at temperatures about 68-72°C. The method uses several restriction enzymes to generate a suitable fragment in the known region of a gene. The fragment is then circularized by intramolecular ligation and used as a PCR template.

Capture PCR (Lagerstrom, M. *et al.*, *PCR Methods Applic.* 1:111-19, (1991)) is a method for PCR amplification of DNA fragments adjacent to a known sequence in human and yeast artificial chromosome DNA. Capture PCR

- 19 -

also requires multiple restriction enzyme digestions and ligations to place an engineered double-stranded sequence into a flanking part of the DNA molecule before PCR.

Another method which may be used to retrieve flanking sequences is that of Parker, J.D., *et al.*, *Nucleic Acids Res.*, **19**:3055-60, (1991)). Additionally, one can use PCR, nested primers and PromoterFinder™ libraries to "walk in" genomic DNA (PromoterFinder™; Clontech, Palo Alto, CA). This process avoids the need to screen libraries and is useful in finding intron/exon junctions. Preferred libraries for screening for full length cDNAs are ones that have been size-selected to include larger cDNAs. Also, random primed libraries are preferred in that they will contain more sequences which contain the 5' and upstream regions of genes.

A randomly primed library may be particularly useful if an oligo d(T) library does not yield a full-length cDNA. Genomic libraries are useful for extension into the 5' nontranslated regulatory region.

The nucleic acid sequences and oligonucleotides of the invention can also be prepared by solid-phase methods, according to known synthetic methods. Typically, fragments of up to about 100 bases are individually synthesized, then joined to form continuous sequences up to several hundred bases.

B. Use of StAR-B nucleic acid sequence for the production of StAR-B products

In accordance with the present invention, nucleic acid sequences specified above may be used as recombinant DNA molecules that direct the expression of StAR-B products.

As will be understood by those of skill in the art, it may be advantageous to produce StAR-B product-encoding nucleotide sequences possessing codons other than those which appear in SEQ ID NO:1 which are those which naturally occur in the human genome. Codons preferred by a particular prokaryotic or

- 20 -

eukaryotic host (Murray, E. *et al. Nuc Acids Res.*, 17:477-508, (1989)) can be selected, for example, to increase the rate of StAR-B product expression or to produce recombinant RNA transcripts having desirable properties, such as a longer half-life, than transcripts produced from naturally occurring sequence.

5 The nucleic acid sequences of the present invention can be engineered in order to alter a StAR-B product coding sequence for a variety of reasons, including but not limited to, alterations which modify the cloning, processing and/or expression of the product. For example, alterations may be introduced using techniques which are well known in the art, e.g., site-directed mutagenesis,
10 to insert new restriction sites, to alter glycosylation patterns, to change codon preference, to produce splice variants, etc.

The present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a nucleic acid
15 sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. Appropriate cloning and
20 expression vectors for use with prokaryotic and eukaryotic hosts are also described in Sambrook, *et al.*, (*supra*).

The present invention also relates to host cells which are genetically engineered with vectors of the invention, and the production of the product of the invention by recombinant techniques. Host cells are genetically engineered (i.e.,
25 transduced, transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as

- 21 -

appropriate for activating promoters, selecting transformants or amplifying the expression of the StAR-B nucleic acid sequence. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to those skilled in the art.

5 The nucleic acid sequences of the present invention may be included in any one of a variety of expression vectors for expressing a product. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral
10 DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host. The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures
15 and related sub-cloning procedures are deemed to be within the scope of those skilled in the art.

The DNA sequence in the expression vector is operatively linked to an appropriate transcription control sequence (promoter) to direct mRNA synthesis. Examples of such promoters include: LTR or SV40 promoter, the *E.coli lac* or
20 *trp* promoter, the phage lambda *PL* promoter, and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation, and a transcription terminator. The vector may also include appropriate sequences for amplifying expression. In addition, the expression
25 vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in *E.coli*.

- 22 -

The vector containing the appropriate DNA sequence as described above, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the protein. Examples of appropriate expression hosts include: bacterial cells, such as *E.coli*,
5 *Streptomyces*, *Salmonella typhimurium*; fungal cells, such as yeast; insect cells such as *Drosophila* and *Spodoptera Sf9*; animal cells such as CHO, COS, HEK 293 or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein. The invention is not limited by the host cells employed.

10 In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the StAR-B product. For example, when large quantities of StAR-B product are needed for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be desirable. Such vectors include, but are not limited to,
15 multifunctional *E.coli* cloning and expression vectors such as *Bluescript*(R) (Stratagene), in which the StAR-B polypeptide coding sequence may be ligated into the vector in-frame with sequences for the amino-terminal Met and the subsequent 7 residues of beta-galactosidase so that a hybrid protein is produced; *pIN* vectors (Van Heeke & Schuster *J. Biol. Chem.* 264:5503-5509, (1989)); *pET*
20 vectors (Novagen, Madison WI); and the like.

In the yeast *Saccharomyces cerevisiae* a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and PGH may be used. For reviews, see Ausubel *et al.* (*supra*) and Grant *et al.*, (*Methods in Enzymology* 153:516-544, (1987)).

25 In cases where plant expression vectors are used, the expression of a sequence encoding StAR-B product may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of *CaMV* (Brisson *et al.*, *Nature* 310:511-514. (1984)) may be used alone or in

combination with the omega leader sequence from TMV (Takamatsu *et al.*, *EMBO J.*, 6:307-311, (1987)). Alternatively, plant promoters such as the small subunit of RUBISCO (Coruzzi *et al.*, *EMBO J.* 3:1671-1680, (1984); Broglie *et al.*, *Science* 224:838-843, (1984)); or heat shock promoters (Winter J and Sinibaldi R.M., *Results Probl. Cell Differ.*, 17:85-105, (1991)) may be used. These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. For reviews of such techniques, see Hobbs S. or Murry L.E. (1992) in McGraw Hill Yearbook of Science and Technology, McGraw Hill, New York, N.Y., pp 191-196; or Weissbach and Weissbach (1988) *Methods for Plant Molecular Biology*, Academic Press, New York, N.Y., pp 421-463.

StAR-B product may also be expressed in an insect system. In one such system, Autographa californica nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The StAR-B product coding sequence may be cloned into a nonessential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of StAR-B coding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein coat. The recombinant viruses are then used to infect *S. frugiperda* cells or *Trichoplusia* larvae in which StAR-B protein is expressed (Smith *et al.*, *J. Virol.* 46:584, (1983); Engelhard, E.K. *et al.*, *Proc. Nat. Acad. Sci.* 91:3224-7, (1994)).

In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, a StAR-B product coding sequence may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a nonessential E1 or E3 region of the viral genome will result in a viable virus capable of expressing StAR-B protein in infected host

- 24 -

cells (Logan and Shenk, *Proc. Natl. Acad. Sci.* 81:3655-59, (1984). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be required for efficient translation of
5 a StAR-B protein coding sequence. These signals include the ATG initiation codon and adjacent sequences. In cases where StAR-B product coding sequence, its initiation codon and upstream sequences are inserted into the appropriate expression vector, no additional translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted,
10 exogenous transcriptional control signals including the ATG initiation codon must be provided. Furthermore, the initiation codon must be in the correct reading frame to ensure transcription of the entire insert. Exogenous transcriptional elements and initiation codons can be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the
15 inclusion of enhancers appropriate to the cell system in use (Scharf, D. *et al.*, (1994) *Results Probl. Cell Differ.*, 20:125-62, (1994); Bittner *et al.*, *Methods in Enzymol* 153:516-544, (1987)).

In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher
20 eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Dibner, M., and Battey, I. (1986) *Basic Methods in Molecular*
25 *Biology*). Cell-free translation systems can also be employed to produce polypeptides using RNAs derived from the DNA constructs of the present invention.

- 25 -

A host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the protein include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation. Post-translational processing which cleaves a "pre-pro" form of the protein may also be important for correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, 293, WI38, etc. have specific cellular machinery and characteristic mechanisms for such post-translational activities and may be chosen to ensure the correct modification and processing of the introduced, foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express StAR-B product may be transformed using expression vectors which contain viral origins of replication or endogenous expression elements and a selectable marker gene. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clumps of stably transformed cells can be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler M., *et al.*, *Cell* 11:223-32, (1977)) and adenine phosphoribosyltransferase (Lowy I., *et al.*, *Cell* 22:817-23, (1980)) genes which can be employed in *tk*- or *aprt*- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, *dhfr* which confers resistance to methotrexate (Wigler M., *et al.*, *Proc. Natl. Acad. Sci.* 77:3567-70, (1980)); *npt*, which confers resistance to the

- 26 -

aminoglycosides neomycin and G-418 (Colbere-Garapin, F. *et al.*, *J. Mol. Biol.*, 150:1-14, (1981)) and *als* or *pat*, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, *trpB*, which allows cells to
5 utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman S.C. and R.C. Mulligan, *Proc. Natl. Acad. Sci.* 85:8047-51, (1988)). The use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate, GUS, and luciferase and its substrates, luciferin and ATP, being widely used not only to
10 identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C.A. *et al.*, *Methods Mol. Biol.*, 55:121-131, (1995)).

Host cells transformed with a nucleotide sequence encoding StAR-B product may be cultured under conditions suitable for the expression and
15 recovery of the encoded protein from cell culture. The product produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing nucleic acid sequences encoding StAR-B product can be designed with signal sequences which direct secretion of StAR-B
20 product through a prokaryotic or eukaryotic cell membrane.

StAR-B product may also be expressed as a recombinant protein with one or more additional polypeptide domains added to facilitate protein purification. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on
25 immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp, Seattle, Wash.). The inclusion of a protease-cleavable polypeptide linker sequence between the purification domain

- 27 -

and StAR-B protein is useful to facilitate purification. One such expression vector provides for expression of a fusion protein comprising a StAR-B polypeptide fused to a polyhistidine region separated by an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography, as described in Porath, *et al.*, *Protein Expression and Purification*, 3:263-281, (1992)) while the enterokinase cleavage site provides a means for isolating StAR-B polypeptide from the fusion protein. *pGEX* vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to ligand-agarose beads (e.g., glutathione-agarose in the case of GST-fusions) followed by elution in the presence of free ligand.

Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, or other methods, which are well known to those skilled in the art.

The StAR-B products can be recovered and purified from recombinant cell cultures by any of a number of methods well known in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography, and lectin chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high

- 28 -

performance liquid chromatography (HPLC) can be employed for final purification steps.

C. Diagnostic applications utilizing nucleic acid sequences

5 The nucleic acid sequences of the present invention may be used for a variety of diagnostic purposes. The nucleic acid sequences may be used to detect and quantitate expression of StAR-B in patient's cells, e.g. biopsied tissues, by detecting the presence of mRNA coding for StAR-B product. Alternatively, the assay may be used to detect soluble StAR-B in the serum or blood. This assay typically involves obtaining total mRNA from the tissue or serum and contacting 10 the mRNA with a nucleic acid probe. The probe is a nucleic acid molecule of at least 20 nucleotides, preferably 20-30 nucleotides, capable of specifically hybridizing with a sequence included within the sequence of a nucleic acid molecule encoding StAR-B under hybridizing conditions, detecting the presence 15 of mRNA hybridized to the probe, and thereby detecting the expression of StAR-B. This assay can be used to distinguish between absence, presence, and excess expression of StAR-B product and to monitor levels of StAR-B expression during therapeutic intervention.

The invention also contemplates the use of the nucleic acid sequences as a 20 diagnostic for diseases resulting from inherited defective StAR-B sequences, or diseases in which the purpose of the amount of the known StAR to the novel StAR variant of the invention is altered. These sequences can be detected by comparing the sequences of the defective (i.e., mutant) StAR-B coding region with that of a normal coding region. Association of the sequence coding for 25 mutant StAR-B product with abnormal StAR-B product activity may be verified. In addition, sequences encoding mutant StAR-B products can be inserted into a suitable vector for expression in a functional assay system (e.g., colorimetric assay, complementation experiments in a StAR-B protein deficient strain of

- 29 -

HEK293 cells) as yet another means to verify or identify mutations. Once mutant genes have been identified, one can then screen populations of interest for carriers of the mutant gene.

Individuals carrying mutations in the nucleic acid sequence of the present invention may be detected at the DNA level by a variety of techniques. Nucleic acids used for diagnosis may be obtained from a patient's cells, including but not limited to such as from blood, urine, saliva, placenta, tissue biopsy and autopsy material. Genomic DNA may be used directly for detection or may be amplified enzymatically by using PCR (Saiki, *et al.*, *Nature* **324**:163-166, (1986)) prior to analysis. RNA or cDNA may also be used for the same purpose. As an example, PCR primers complementary to the nucleic acid of the present invention can be used to identify and analyze mutations in the gene of the present invention. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype.

Point mutations can be identified by hybridizing amplified DNA to radiolabeled RNA of the invention or alternatively, radiolabeled antisense DNA sequences of the invention. Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (e.g. Cotton, *et al.* *Proc. Natl. Acad. Sci. USA*, **85**:4397-4401, (1985)), or by differences in melting temperatures. "Molecular beacons" (Kostrikis L.G. *et al.*, *Science* **279**:1228-1229, (1998)), hairpin-shaped, single-stranded synthetic oligo-nucleotides containing probe sequences which are complementary to the nucleic acid of the present invention, may also be used to detect point mutations or other sequence changes as well as monitor expression levels of StAR-B product. Such diagnostics would be particularly useful for prenatal testing.

Another method for detecting mutations uses two DNA probes which are designed to hybridize to adjacent regions of a target, with abutting bases, where

- 30 -

the region of known or suspected mutation(s) is at or near the abutting bases. The two probes may be joined at the abutting bases, e.g., in the presence of a ligase enzyme, but only if both probes are correctly base paired in the region of probe junction. The presence or absence of mutations is then detectable by the
5 presence or absence of ligated probe.

Also suitable for detecting mutations in the StAR-B product coding sequence are oligonucleotide array methods based on sequencing by hybridization (SBH), as described, for example, in U.S. Patent No. 5,547,839. In a typical method, the DNA target analyte is hybridized with an array of
10 oligonucleotides formed on a microchip. The sequence of the target can then be "read" from the pattern of target binding to the array.

D. Gene mapping utilizing nucleic acid sequences

The nucleic acid sequences of the present invention are also valuable for
15 chromosome identification. The sequence is specifically targeted to and can hybridize with a particular location on an individual human chromosome. Moreover, there is a current need for identifying particular sites on the chromosome. Few chromosome marking reagents based on actual sequence data (repeat polymorphisms) are presently available for marking chromosomal
20 location. The mapping of DNAs to chromosomes according to the present invention is an important first step in correlating those sequences with genes associated with disease.

Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 20-30 bp) from the StAR-B cDNA. Computer analysis of the
25 3' untranslated region is used to rapidly select primers that do not span more than one exon in the genomic DNA, which would complicate the amplification process. These primers are then used for PCR screening of somatic cell hybrids

- 31 -

containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the primer will yield an amplified fragment.

PCR mapping of somatic cell hybrids or using instead radiation hybrids are rapid procedures for assigning a particular DNA to a particular chromosome.

5 Using the present invention with the same oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes or pools of large genomic clones in an analogous manner. Other mapping strategies that can similarly be used to map to its chromosome include *in situ* hybridization, prescreening with labeled flow-sorted chromosomes and
10 preselection by hybridization to construct chromosome specific-cDNA libraries.

Fluorescence *in situ* hybridization (FISH) of a cDNA clone to a metaphase chromosomal spread can be used to provide a precise chromosomal location in one step. This technique can be used with cDNA as short as 50 or 60 bases. For a review of this technique, see Verma *et al.*, *Human Chromosomes: a Manual of*
15 *Basic Techniques*, (1988) Pergamon Press, New York.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found, for example, in the OMIM database (Center for Medical Genetics, Johns Hopkins University, Baltimore, MD and
20 National Center for Biotechnology Information, National Library of Medicine, Bethesda, MD). The OMIM gene map presents the cytogenetic map location of disease genes and other expressed genes. The OMIM database provides information on diseases associated with the chromosomal location. Such associations include the results of linkage analysis mapped to this interval, and
25 the correlation of translocations and other chromosomal aberrations in this area with the advent of polygenic diseases, such as cancer, in general and prostate cancer in particular.

E. Therapeutic applications of nucleic acid sequences

Nucleic acid sequences of the invention may also be used for therapeutic purposes. Turning first to the second aspect of the invention (i.e. inhibition of expression of StAR-B), expression of StAR-B product may be modulated through
5 antisense technology, which controls gene expression through hybridization of complementary nucleic acid sequences, i.e. antisense DNA or RNA, to the control, 5' or regulatory regions of the gene encoding StAR-B product. For example, the 5' coding portion of the nucleic acid sequence sequence which codes for the product of the present invention is used to design an antisense
10 oligonucleotide of from about 10 to 40 base pairs in length. Oligonucleotides derived from the transcription Start site, e.g. between positions -10 and +10 from the Start site, are preferred. An antisense DNA oligonucleotide is designed to be complementary to a region of the nucleic acid sequence involved in transcription (Lee *et al.*, *Nucl. Acids, Res.*, 6:3073, (1979); Cooney *et al.*, *Science* 241:456,
15 (1988); and Dervan *et al.*, *Science* 251:1360, (1991)), thereby preventing transcription and the production of the StAR-B products. An antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the StAR-B products (Okano *J. Neurochem.* 56:560, (1991)). The antisense constructs can be delivered to cells by procedures known
20 in the art such that the antisense RNA or DNA may be expressed *in vivo*. The antisense may be antisense mRNA or DNA sequence capable of coding such antisense mRNA. The antisense mRNA or the DNA coding thereof can be complementary to the full sequence of nucleic acid sequences coding to the StAR-B protein or to a fragment of such a sequence which is sufficient to inhibit
25 production of a protein product.

Turning now to the first aspect of the invention, i.e. expression of StAR-B, expression of StAR-B product may be increased by providing coding sequences

- 33 -

for coding for said product under the control of suitable control elements ending its expression in the desired host.

The nucleic acid sequences of the invention may be employed in combination with a suitable pharmaceutical carrier. Such compositions comprise
5 a therapeutically effective amount of the compound, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The formulation should suit the mode of administration.

The products of the invention as well as any activators and deactivators
10 compounds (see below) which are polypeptides, may also be employed in accordance with the present invention by expression of such polypeptides *in vivo*, which is often referred to as "*gene therapy*." Cells from a patient may be engineered with a nucleic acid sequence (DNA or RNA) encoding a polypeptide *ex vivo*, with the engineered cells then being provided to a patient to be treated
15 with the polypeptide. Such methods are well-known in the art. For example, cells may be engineered by procedures known in the art by use of a retroviral particle containing RNA encoding a polypeptide of the present invention.

Similarly, cells may be engineered *in vivo* for expression of a polypeptide *in vivo* by procedures known in the art. As known in the art, a producer cell for
20 producing a retroviral particle containing RNA encoding the polypeptide of the present invention may be administered to a patient for engineering cells *in vivo* and expression of the polypeptide *in vivo*. These and other methods for administering a product of the present invention by such method should be apparent to those skilled in the art from the teachings of the present invention.
25 For example, the expression vehicle for engineering cells may be other than a retrovirus, for example, an adenovirus which may be used to engineer cells *in vivo* after combination with a suitable delivery vehicle.

- 34 -

Retroviruses from which the retroviral plasmid vectors mentioned above may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus, retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, adenovirus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the *PE501*, *PA317*, *psi-2*, *psi-AM*, *PA12*, *T19-14X*, *VT-19-17-H2*, *psi-CRE*, *psi-CRIP*, *GP+E-86*, *GP+envAm12*, and *DAN* cell lines as described in Miller (*Human Gene Therapy*, Vol. 1, pg. 5-14, (1990)). The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and CaPO_4 precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

The producer cell line generates infectious retroviral vector particles which include the nucleic acid sequence(s) encoding the polypeptides. Such retroviral vector particles then may be employed, to transduce eukaryotic cells, either in vitro or in vivo. The transduced eukaryotic cells will express the nucleic acid sequence(s) encoding the polypeptide. Eukaryotic cells which may be transduced include, but are not limited to, embryonic stem cells, embryonic carcinoma cells, as well as hematopoietic stem cells, hepatocytes, fibroblasts, myoblasts, keratinocytes, endothelial cells, and bronchial epithelial cells.

The genes introduced into cells may be placed under the control of inducible promoters, such as the radiation-inducible Egr-1 promoter, (Maceri, H.J., *et al.*, *Cancer Res.*, 56(19):4311 (1996)), to stimulate StAR-B production or antisense inhibition in response to radiation, eg., radiation therapy for treating tumors.

Example II. StAR-B product

The substantially purified StAR-B product of the invention has been defined above as the product coded from the nucleic acid sequence of the invention. Preferably the amino acid sequence is an amino acid sequence having at least 70%, preferably at least 80% or 90% identity to the sequence identified as SEQ ID NO:2. The protein or polypeptide may be in mature and/or modified form, also as defined above. Also contemplated are protein fragments having at least 10 contiguous amino acid residues, preferably at least 10-20 residues, derived from the StAR-B product.

The sequence variations are preferably those that are considered conserved substitutions, as defined above. Thus, for example, a protein with a sequence having at least 80% sequence identity with the protein identified as SEQ ID NO:2, preferably by utilizing conserved substitutions as defined above is also part of the invention. In a more specific embodiment, the protein has or contains the sequence identified SEQ ID NO:2. The StAR-B product may be (i) one in which one or more of the amino acid residues in a sequence listed above are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue), or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the StAR-B product is fused with another compound, such as a compound to increase the half-life of the protein (for example, polyethylene glycol (PEG)), or a moiety which serves as targeting means to direct the protein to its target tissue or target cell population (such as an antibody), or (iv) one in which additional amino acids are fused to the StAR-B product. Such fragments, variants and derivatives are deemed to be within the scope of those skilled in the art from the teachings herein.

A. Preparation of StAR-B product

Recombinant methods for producing and isolating the StAR-B product, and fragments of the protein are described above.

In addition to recombinant production, fragments and portions of StAR-B product may be produced by direct peptide synthesis using solid-phase techniques (cf. Stewart *et al.*, (1969) Solid-Phase Peptide Synthesis, WH Freeman Co, San Francisco; Merrifield J., *J. Am. Chem. Soc.*, 85:2149-2154, (1963)). In vitro peptide synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer, Foster City, Calif.) in accordance with the instructions provided by the manufacturer. Fragments of StAR-B product may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

B. Therapeutic uses and compositions utilizing the StAR-B product

The StAR-B product of the invention is generally useful in treating diseases and disorders which are characterized by a lower than normal level of StAR-B expression, and or diseases which can be cured or ameliorated by raising the level of the StAR-B product, even if the level is normal.

StAR-B products or fragments may be administered by any of a number of routes and methods designed to provide a consistent and predictable concentration of compound at the target organ or tissue. The product-containing compositions may be administered alone or in combination with other agents, such as stabilizing compounds, and/or in combination with other pharmaceutical agents such as drugs or hormones.

StAR-B product-containing compositions may be administered by a number of routes including, but not limited to oral, intravenous, intramuscular, transdermal, subcutaneous, topical, sublingual, or rectal means as well as by nasal application. StAR-B product-containing compositions may also be administered

- 37 -

via liposomes. Such administration routes and appropriate formulations are generally known to those of skill in the art.

The product can be given via intravenous or intraperitoneal injection. Similarly, the product may be injected to other localized regions of the body. The product may also be administered via nasal insufflation. Enteral administration is also possible. For such administration, the product should be formulated into an appropriate capsule or elixir for oral administration, or into a suppository for rectal administration.

The foregoing exemplary administration modes will likely require that the product be formulated into an appropriate carrier, including ointments, gels, suppositories. Appropriate formulations are well known to persons skilled in the art.

Dosage of the product will vary, depending upon the potency and therapeutic index of the particular polypeptide selected.

A therapeutic composition for use in the treatment method can include the product in a sterile injectable solution; the polypeptide in an oral delivery vehicle, the product in an aerosol suitable for nasal administration, or the product in a nebulized form, all prepared according to well known methods. Such compositions comprise a therapeutically effective amount of the compound, and a pharmaceutically acceptable carrier or excipient. Such a carrier includes but is not limited to saline, buffered saline, dextrose, water, glycerol, ethanol, and combinations thereof. The product of the invention may also be used to modulate endothelial differentiation and proliferation as well as to modulate apoptosis either *ex vivo* or *in vitro*, for example, in cell cultures.

Example III. Screening methods for activators and deactivators (inhibitors)

The present invention also includes an assay for identifying molecules, such as synthetic drugs, antibodies, peptides, or other molecules, which have a

modulating effect on the activity of the StAR-B product, e.g. activators or deactivators of the StAR-B product of the present invention. Such an assay comprises the steps of providing an StAR-B product encoded by the nucleic acid sequences of the present invention, contacting the StAR-B protein with one or
5 more candidate molecules to determine the candidate molecules modulating effect on the activity of the StAR-B product, and selecting from the molecules a candidate's molecule capable of modulating StAR-B product physiological activity.

StAR-B product, its catalytic or immunogenic fragments or oligopeptides
10 thereof, can be used for screening therapeutic compounds in any of a variety of drug screening techniques. The fragment employed in such a test may be free in solution, affixed to a solid support, borne on a cell membrane or located intracellularly. The formation of binding complexes, between StAR-B product and the agent being tested, may be measured. Alternatively, the activator or
15 deactivator may work by serving as agonist or antagonist, respectively, of the StAR-B receptor and their effect may be determined in connection with the receptor.

Another technique for drug screening which may be used provides for high throughput screening of compounds having suitable binding affinity to the
20 StAR-B product is described in detail by Geysen in PCT Application WO 84/03564, published on Sep. 13, 1984. In summary, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with the full StAR-B product or with fragments of StAR-B product and washed. Bound
25 StAR-B product is then detected by methods well known in the art. Substantially purified StAR-B product can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing

- 39 -

antibodies can be used to capture the peptide and immobilize it on a solid support.

Antibodies to the StAR-B product, as described in Example IV below, may also be used in screening assays according to methods well known in the art.

5 For example, a "*sandwich*" assay may be performed, in which an anti-StAR-B antibody is affixed to a solid surface such as a microtiter plate and StAR-B product is added. Such an assay can be used to capture compounds which bind to the StAR-B product. Alternatively, such an assay may be used to measure the ability of compounds to influence with the binding of StAR-B product to the
10 StAR-B receptor [I and then select those compounds which effect the binding.

Example IV. Anti-StAR-B antibodies

A. Synthesis

In still another aspect of the invention, the purified StAR-B product is
15 used to produce anti-StAR-B antibodies which have diagnostic and therapeutic uses related to the activity, distribution, and expression of the StAR-B product, in particular therapeutic applications in inhibiting the effect of the StAR-B cholesterol in the transport.

Antibodies to StAR-B product may be generated by methods well known
20 in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, humanized, single chain, Fab fragments and fragments produced by an Fab expression library. Antibodies, i.e., those which inhibit dimer formation, are especially preferred for therapeutic use.

A fragment StAR-B product for antibody induction does not require
25 biological activity but have to feature immunological activity; however, the protein fragment or oligopeptide must be antigenic. Peptides used to induce specific antibodies may have an amino acid sequence consisting of at least five amino acids, preferably at least 10 amino acids of the sequences specified in SEQ

ID NO: 2. Preferably they should mimic a portion of the amino acid sequence of the natural protein and may contain the entire amino acid sequence of a small, naturally occurring molecule. Short stretches of StAR-B protein amino acids may be fused with those of another protein such as keyhole limpet hemocyanin and
5 antibody produced against the chimeric molecule. Procedures well known in the art can be used for the production of antibodies to StAR-B product.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, etc may be immunized by injection with StAR-B product or any portion, fragment or oligopeptide which retains immunogenic properties.
10 Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include but are not limited to Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, and dinitrophenol. BCG (bacilli Calmette-Guerin) and
15 *Corynebacterium parvum* are potentially useful human adjuvants.

Monoclonal antibodies to StAR-B protein may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include but are not limited to the hybridoma technique originally described by Koehler and Milstein (*Nature* 256:495-497, (1975)), the
20 human B-cell hybridoma technique (Kosbor *et al.*, *Immunol. Today* 4:72, (1983); Cote *et al.*, *Proc. Natl. Acad. Sci.* 80:2026-2030, (1983)) and the EBV-hybridoma technique (Cole, *et al.*, *Mol. Cell Biol.* 62:109-120, (1984)).

Techniques developed for the production of "*chimeric antibodies*", the splicing of mouse antibody genes to human antibody genes to obtain a molecule
25 with appropriate antigen specificity and biological activity can also be used (Morrison *et al.*, *Proc. Natl. Acad. Sci.* 81:6851-6855, (1984); Neuberger *et al.*, *Nature* 312:604-608, (1984); Takeda *et al.*, *Nature* 314:452-454, (1985)). Alternatively, techniques described for the production of single chain antibodies

- 41 -

(U.S. Pat. No. 4,946,778) can be adapted to produce single-chain antibodies specific for the StAR-B protein.

Antibodies may also be produced by inducing *in vivo* production in the lymphocyte population or by screening recombinant immunoglobulin libraries or
5 panels of highly specific binding reagents as disclosed in Orlandi *et al.* (*Proc. Natl. Acad. Sci.* **86**:3833-3837, 1989)), and Winter G and Milstein C., (*Nature* **349**:293-299, (1991)).

Antibody fragments which contain specific binding sites for StAR-B protein may also be generated. For example, such fragments include, but are not
10 limited to, the F(ab')₂ fragments which can be produced by pepsin digestion of the antibody molecule and the Fab fragments which can be generated by reducing the disulfide bridges of the F(ab')₂ fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity (Huse W.D. *et al.*, *Science*
15 256:1275-1281, (1989)).

B. Diagnostic applications of antibodies

A variety of protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established
20 specificities are well known in the art. Such immunoassays typically involve the formation of complexes between StAR-B product and its specific antibody and the measurement of complex formation. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two noninterfering epitopes on a specific StAR-B product is preferred, but a competitive binding
25 assay may also be employed. These assays are described in Maddox D.E., *et al.*, (*J. Exp. Med.* **158**:1211, (1983)).

Antibodies which specifically bind StAR-B product are useful for the diagnosis of conditions or diseases characterized by over or under expression of

- 42 -

StAR-B. Alternatively, such antibodies may be used in assays to monitor patients being treated with StAR-B product, its activators, or its deactivators. Diagnostic assays for StAR-B protein include methods utilizing the antibody and a label to detect StAR-B product in human body fluids or extracts of cells or tissues. The products and antibodies of the present invention may be used with or without modification. Frequently, the proteins and antibodies will be labeled by joining them, either covalently or noncovalently, with a reporter molecule. A wide variety of reporter molecules are known in the art.

A variety of protocols for measuring StAR-B product, using either polyclonal or monoclonal antibodies specific for the respective protein are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescent activated cell sorting (FACS). As noted above, a two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on StAR-B product is preferred, but a competitive binding assay may be employed. These assays are described, among other places, in Maddox, *et al. (supra)*. Such protocols provide a basis for diagnosing altered or abnormal levels of StAR-B product expression. Normal or standard values for StAR-B product expression are established by combining body or cell extracts taken from normal subjects, preferably human, with antibody to StAR-B product under conditions suitable for complex formation which are well known in the art. The amount of standard complex formation may be quantified by various methods, preferably by photometric methods. Then, standard values obtained from normal samples may be compared with values obtained from samples from subjects potentially affected by disease. Deviation between standard and subject values establishes the presence of disease state.

The antibody assays are useful to determine the level of StAR-B present in a body fluid sample, in order to determine whether it is being overexpressed or

underexpressed in the tissue, or as an indication of how StAR-B levels are responding to drug treatment.

C. Therapeutic uses of antibodies

5 In addition to their diagnostic use the antibodies may have a therapeutical utility in blocking or decreasing the activity of the StAR-B product in pathological conditions where beneficial effect can be achieved by such a decrease.

The antibody employed is preferably a humanized monoclonal antibody,
10 or a human Mab produced by known globulin-gene library methods. The antibody is administered typically as a sterile solution by IV injection, although other parenteral routes may be suitable. Typically, the antibody is administered in an amount between about 1-15 mg/kg body weight of the subject. Treatment is continued, e.g., with dosing every 1-7 days, until a therapeutic improvement is
15 seen.

Although the invention has been described with reference to specific methods and embodiments, it is appreciated that various modifications and changes may be made without departing from the invention.

- 44 -

SEQUENCE LISTING

<110> COMPUGEN LTD.

5 <120> NOVEL AMINO ACID AND NUCLEIC ACID SEQUENCES

<130> COMPU1

<140>

10 <141>

<160> 2

<170> PatentIn Ver. 2.0

15 <210> 1

<211> 1615

<212> DNA

<213> human

20 <400> 1

```

ccgcagctaa ggcactcgc cgacgcaatg gacccggcgc tggcagccca gatgagcgag 60
gctgtggcca gaagatgctc cagtaccggc gggacacagc aggctggaag atttgccggg 120
aaggcaangg agtttcagtt tcctggaggc catctgtgga gtttccaggg aacctgtacc 180
25 gaggagaagg cattgtatat gggacactag aggaggtgtg ggactgtgtg aagccagctg 240
ttggaggcct acgagtgaag tgggatgaga atgtgaccgg ttttgaaatt atccaaagca 300
tcactgacac cctgtgtgta agcagaacct ccactccctc cgctgccatg aagctcattt 360
ctcccagaga ttttgtggac ttgtgtgtag tcaagagata tgaggatggg accatcagtt 420
ccaacgccac ccatgtggag catccgttat gtcccccga ggcaggtttt gtgagaggat 480
30 ttaaccatcc ttgtggttgc ttctgtgaac ctcttccagg ggaacccacc aagaccaacc 540
tggtcacatt cttccatacc gacctcagcg gttacctccc acagaacgtg gtggactcct 600
tcttcccccg cagcatgacc cggttttatg ccaaccttca gaaagcagtg aagcaattcc 660
atgagtaatg ctatcgttac ttcttggcaa agaactcccg tgactcatcg aggagctcca 720
gctgttggga caccaaggag cctgggagca cgcagaggcc tgtgttcact ctttggaaaca 780
35 agctgatgga ctgcgcatct ctgagaatgc caaccagagg cggcagccca cccttcctgc 840
ctcctgcccc actcagggtt ggcgtgtgat gagccattca tgtgttccaa actccatctg 900
cctgttacct aaacacgcct ctctggcag ggtagaccca ggcctctaac catctgacag 960
agactcggcc tggacaccat gcgatgcact ctggcaccaa ggctttatgt gccatcact 1020
ctcagagacc acgtttccct gactgtcata gagaatcatc atcgccactg aaaaccaggc 1080
40 cctgttgccct ttaagcatg taccgctccc tcagtcctgt gctgcagccc cccaaatata 1140
tttttctgat atagaccttg tatatggctt taatgccgca aaatatttat ttttccttaa 1200
aaaagggtgc aacttggaaa taatggttta aaaacaggat aagcattaag gaaaaacact 1260
ttcaatgtgt cttccatttg atgaatttgt ttttctctct ttatccccgc aagtggagtt 1320
tcattgtcctc ggtgaaacca gacagtgtga atctgttcca gcccaaactc gcagcattag 1380
45 ggatgagttc tcrgaagtga ttctgaactg agcacgcact catgtctgca tggggaaactc 1440
tggggagaag agccttcctt ttctttccct tgggccattt gcctttcctt gtcgtcttac 1500
tgagggcgga ggcaggagg gtctctgtct ttccagggcc ctgggcaggg ccatcctggc 1560
cattcaggga aagatgggaa gagttagggc tccgttttag gcagcctggg tggga 1615

```

50 <210> 2

<211> 221

<212> PRT

<213> human

55 <400> 2

Ala Ala Lys Arg Thr Arg Arg Arg Asn Gly Pro Gly Ala Gly Ser Pro
1 5 10 15

Asp Glu Arg Gly Cys Gly Gln Lys Met Leu Gln Tyr Arg Arg Asp Thr
60 20 25 30

Ala Gly Trp Lys Ile Cys Arg Glu Gly Xaa Gly Val Ser Val Ser Trp
35 40 45

- 45 -

Arg Pro Ser Val Glu Phe Pro Gly Asn Leu Tyr Arg Gly Glu Gly Ile
 50 55 60
 5 Val Tyr Gly Thr Leu Glu Glu Val Trp Asp Cys Val Lys Pro Ala Val
 65 70 75 80
 Gly Gly Leu Arg Val Lys Trp Asp Glu Asn Val Thr Gly Phe Glu Ile
 85 90 95
 10 Ile Gln Ser Ile Thr Asp Thr Leu Cys Val Ser Arg Thr Ser Thr Pro
 100 105 110
 15 Ser Ala Ala Met Lys Leu Ile Ser Pro Arg Asp Phe Val Asp Leu Val
 115 120 125
 Leu Val Lys Arg Tyr Glu Asp Gly Thr Ile Ser Ser Asn Ala Thr His
 130 135 140
 20 Val Glu His Pro Leu Cys Pro Pro Lys Pro Gly Phe Val Arg Gly Phe
 145 150 155 160
 Asn His Pro Cys Gly Cys Phe Cys Glu Pro Leu Pro Gly Glu Pro Thr
 165 170 175
 25 Lys Thr Asn Leu Val Thr Phe Phe His Thr Asp Leu Ser Gly Tyr Leu
 180 185 190
 30 Pro Gln Asn Val Val Asp Ser Phe Phe Pro Arg Ser Met Thr Arg Phe
 195 200 205
 Tyr Ala Asn Leu Gln Lys Ala Val Lys Gln Phe His Glu
 210 215 220

CLAIMS

1. An isolated nucleic acid sequence selected from the group consisting of:
 - (i) the nucleic acid sequence depicted in SEQ ID NO: 1;
 - (ii) nucleic acid sequences having at least 70% identity with the
5 sequence of (i); and
 - (iii) fragments of (i) or (ii) of at least 20 b.p.
2. A nucleic acid sequence according to Claim 1(ii) wherein the nucleic acid sequences have at least 80% identity with the sequence of Claim 1(i).
3. A nucleic acid sequence according to Claim 2, wherein the nucleic acid
10 sequences have at least 90% identity.
4. An isolated nucleic acid sequence complementary to the nucleic acid sequence of Claim 1.
5. An amino acid sequence selected from the group consisting of:
 - (i) an amino acid sequence coded by the isolated nucleic acid sequence
15 of Claim 1;
 - (ii) fragments of the amino acid sequence of (i) having at least 10 amino acids;
 - (iii) analogues of the amino acid sequences of (i) or (ii) in which one or more amino acids has been added, deleted, replaced or chemically modified
20 without substantially altering the biological activity of the parent amino acid sequence.
6. An amino acid sequence according to Claim 5, as depicted in SEQ ID NO:2.
7. An isolated nucleic acid sequence coding for the amino acid sequence of Claim 5.
- 25 8. A purified antibody which binds specifically to the amino acid sequence of Claim 5.
9. An expression vector comprising the nucleic acid sequences of Claim 1 and control elements for the expression of the nucleic acid sequence in a suitable host.

10. An expression vector comprising the nucleic acid sequence of Claim 4, and control elements for the expression of the nucleic acid sequence in a suitable host.
11. A host cell transfected by the expression vector of Claim 9.
12. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and as an active ingredient an agent selected from the group consisting of:
- 5 (i) the expression vector of Claim 9; and
- (ii) the amino acid sequence of Claim 5.
13. A pharmaceutical composition according to Claim 12, for treatment of diseases which can be ameliorated or cured by raising the level of the steroidogenic
- 10 acute regulatory protein homolog B (StAR-B).
14. A pharmaceutical composition comprising a pharmaceutically acceptable carrier and as an active ingredient an agent selected from the group consisting of:
- (i) the nucleic acid sequence of Claim 4;
- (ii) the expression vector of Claim 10; and
- 15 (iii) the purified antibody of Claim 8.
15. A pharmaceutical composition according to Claim 14, for treatment of diseases which can be ameliorated or cured by decreasing the level of the StAR-B product.
16. A pharmaceutical composition according to Claim 12, for regulating the
- 20 levels of steroids
17. A pharmaceutical composition according to Claim 16 for regulating the levels of androgen.
18. A method for detecting an StAR-B nucleic acid sequence in a biological sample, comprising the steps of:
- 25 (a) hybridizing to nucleic acid material of said biological sample a nucleic acid sequence of Claim 1; and
- (b) detecting said hybridization complex;
- wherein the presence of said hybridization complex correlates with the presence of an StAR-B nucleic acid sequence in the said biological sample.

19. A method according to Claim 18, wherein the nucleic acid material of said biological sample are mRNA transcripts.

20. A method according to Claim 18, where the nucleic acid sequence is present in a nucleic acid chip.

5 21. A method for identifying candidate compounds capable of binding to the StAR-B product and modulating its activity the method comprising:

(i) providing a protein or polypeptide comprising an amino acid sequence substantially as depicted in SEQ ID NO: 2, or a fragment of such a sequence;

10 (ii) contacting a candidate compound with said amino acid sequence;

(iii) determining the effect of said candidate compound on the biological activity of said protein or polypeptide and selecting those compounds which show a significant effect on said biological activity.

15 22. A method according to Claim 21, wherein the compound is an activator and the measured effect is increase in the biological activity.

23. A method according to Claim 21, wherein the compound is an deactivator and the effect is decrease in the biological activity.

24. An activator of the amino acid sequence of Claim 5.

25. An deactivator of the amino acid sequence of Claims 5.

20 26. A method for detecting SrAR-B-product in a biological sample, comprising the steps of:

(a) contacting with said biological sample the antibody of Claim 8, thereby forming an antibody-antigen complex; and

(b) detecting said antibody-antigen complex

25 wherein the presence of said antibody-antigen complex correlates with the presence of StAR-B product in said biological sample.

1/4

```

ccgcagctaa gcgcactcgc cgacgcaatg gaccgcgcgc tggcagccca gatgagcgag 60
gctgtggcca gaagatgctc cagtaccggc gggacacagc aggctggaag atttgccggg 120
aaggcaangg agtttcagtt tcctggaggc catctgtgga gtttccaggg aacctgtacc 180
gaggagaagg cattgtatat gggacactag aggaggtgtg ggactgtgtg aagccagctg 240
ttggaggcct acgagtgaag tgggatgaga atgtgaccgg ttttgaaatt atccaaagca 300
tcactgacac cctgtgtgta agcagaacct ccactccctc cgctgccatg aagctcattt 360
ctcccagaga ttttgtggac ttggtgctag tcaagagata tgaggatggg accatcagtt 420
ccaacgccac ccatgtggag catccgttat gtccccgaa gccaggtttt gtgagaggat 480
ttaaccatcc ttgtggttgc ttctgtgaac ctcttccagg ggaaccacc aagaccaacc 540
tggtcacatt ctccataacc gacctcagcg gttacctccc acagaacgtg gtggactcct 600
tcttcccccg cagcatgacc cggttttatg ccaaccttca gaaagcagtg aagcaattcc 660
atgagtaatg ctatcggttac ttcttggcaa agaactcccg tgactcatcg aggagctcca 720
gctgttgga caccaaggag cctgggagca cgcagaggcc tgtgttcact ctttggaaca 780
agctgatgga ctgcgcatct ctgagaatgc caaccagagg cggcagccca cccttcctgc 840
ctcctgcccc actcagggtt ggcgtgtgat gagccattca tgtgttccaa actccatctg 900
cctgttacct aaacacgcct ctctggcag ggtagacca ggcctctaac catctgacag 960
agactcggcc tggacaccat gcgatgcact ctggcaccaa ggctttatgt gccatcact 1020
ctcagagacc acgtttccct gactgtcata gagaatcatc atcgccactg aaaaccaggc 1080
cctgttgect ttttaagcatg taccgctccc tcagtcctgt gctgcagccc cccaaatata 1140
tttttctgat atagaccttg tatatggctt taatgccgca aaatatttat ttttccttaa 1200
aaaagggtgc aacttgga aaatgggtta aaaacaggat aagcattaag gaaaaacact 1260
ttcaatgtgt ctccatttg atgaatttg ttttctctct ttatccccgc aagtggagt 1320
tcatgtctc ggtgaaacca gacagtgtga atctgttcca gcccaaact gcagcattag 1380
ggatgagttc tcrgaagtga ttctgaactg agcacgcact catgtctgca tggggaactc 1440
tggggagaag agccttcctt ttcttccct tgggccattt gcctttcctt gtcgtcttac 1500
tgagggcgga ggcaggagg gtctctgtct ttccagggcc ctgggcaggg ccatcctggc 1560
cattcaggga aagatgggaa gagttagggc tccgttttag gcagcctggg tggga 1615

```

Fig. 1

2/4

Ala	Ala	Lys	Arg	Thr	Arg	Arg	Arg	Asn	Gly	Pro	Gly	Ala	Gly	Ser	Pro
1				5					10					15	
Asp	Glu	Arg	Gly	Cys	Gly	Gln	Lys	Met	Leu	Gln	Tyr	Arg	Arg	Asp	Thr
				20					25					30	
Ala	Gly	Trp	Lys	Ile	Cys	Arg	Glu	Gly	Xaa	Gly	Val	Ser	Val	Ser	Trp
				35					40					45	
Arg	Pro	Ser	Val	Glu	Phe	Pro	Gly	Asn	Leu	Tyr	Arg	Gly	Glu	Gly	Ile
				50					55					60	
Val	Tyr	Gly	Thr	Leu	Glu	Glu	Val	Trp	Asp	Cys	Val	Lys	Pro	Ala	Val
				65					70					75	80
Gly	Gly	Leu	Arg	Val	Lys	Trp	Asp	Glu	Asn	Val	Thr	Gly	Phe	Glu	Ile
				85					90					95	
Ile	Gln	Ser	Ile	Thr	Asp	Thr	Leu	Cys	Val	Ser	Arg	Thr	Ser	Thr	Pro
				100					105					110	
Ser	Ala	Ala	Met	Lys	Leu	Ile	Ser	Pro	Arg	Asp	Phe	Val	Asp	Leu	Val
				115					120					125	
Leu	Val	Lys	Arg	Tyr	Glu	Asp	Gly	Thr	Ile	Ser	Ser	Asn	Ala	Thr	His
				130					135					140	

Fig. 2

3/4

Val Glu His Pro Leu Cys Pro Pro Lys Pro Gly Phe Val Arg Gly Phe
145 150 155 160

Asn His Pro Cys Gly Cys Phe Cys Glu Pro Leu Pro Gly Glu Pro Thr
165 170 175

Lys Thr Asn Leu Val Thr Phe Phe His Thr Asp Leu Ser Gly Tyr Leu
180 185 190

Pro Gln Asn Val Val Asp Ser Phe Phe Pro Arg Ser Met Thr Arg Phe
195 200 205

Tyr Ala Asn Leu Gln Lys Ala Val Lys Gln Phe His Glu
210 215 220

Fig. 2 (continued)

4/4

(S79669) steroidogenic acute regulatory protein, StAR [human, follicle cells, Peptide Mitochondrial, 285 aa] [Homo sapiens]gi|2661026 (AF035277) steroidogenic acute regulatory protein [Homo sapiens]

		alignment_scores:
Quality:	169.00	Length: 203
Ratio:	1.420	Gaps: 6
Percent Similarity:	58.621	Percent Identity: 26.108

alignment_block:
cchl10_t53767_0.pep x gi|1172250|bbs|171384 ..

Align seg 1/1 to: gi|1172250|bbs|171384 from: 1 to: 285

```

23 QKMLQYRRDTAGWKICRE.....GXGVSWSWRPSVEFFPGNLYRGEIVYG 67
   || | : ||| :| | : |:| |::| | :|
84 QKALGILSNQEGWK..KESQQDNGDKVMSKVVPDV...GKVFRLEVVDQ 128

68 TLEEVWDCV...KPAVGGLRVKWDENVTFGEIIQSITDLCVSRTSTPSA 114
   :|:::~ :| :| :| :| :| :| :|
129 PMERLYEELVERMEAMG....EWNPNVKEIKVLQKIGKDTFITHELAEEA 174

115 AMKLISPRDFVDLVLVKRYEDGTISSNATHVEHPLCPPKPGFVRGFNHPC 164
   | :|::| ||| : :| :| :| :| :| :|
175 AGNLVGPRDFVSVRCAKR.RGSTCVLAGMATDFGNMPEQKGVIRAEHGPT 223

165 GCFCEPLPGEPKTNLVTFFHTDLGYPQNVVDSFFPRSMTRFYANLQK 214
   :|| |::| || : : ||:| :| :| :| :| :|
224 CMVLHPLAGSPSKTKLTWLLSIDLKGWLPKSIINQVLSQTQVDFANHLRK 273

215 AVK 217
      ::
274 RLE 276

```

Fig. 3

INTERNATIONAL SEARCH REPORT

Int. J. Application No
PCT/IL 00/00252

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/12 C07K14/47 C12Q1/68 A61K38/17 G01N33/50
G01N33/53

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N C07K C12Q A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

STRAND, EPO-Internal, EMBL, WPI Data, PAJ, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>WO 98 20165 A (WHITEHEAD BIOMEDICAL INST ; HUDSON THOMAS (US); LANDER ERIC S (US);) 14 May 1998 (1998-05-14) Note: 100.0 % nt seq identity of SEQ ID NO:3418 (WI-21376b) with SEQ ID NO:1 in 251 bp overlap. the whole document page 14, line 25 -page 28, line 9 page 265, entry WI-21376b --- -/--</p>	<p>1-4, 9-11, 18-20</p>

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"Z" document member of the same patent family

Date of the actual completion of the international search

24 August 2000

Date of mailing of the international search report

06/09/2000

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2260 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

van de Kamp, M

INTERNATIONAL SEARCH REPORT

Int. J. Appl. Application No

PCT/IL 00/00252

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE STRAND 'Online! GenBank; Locus (Accession) AA026809, 9 May 1997 (1997-05-09) HILLIER L ET AL.: "zk02a01.r1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone IMAGE:469320 5', mRNA sequence" XP002145549 Note: 98.3 % nt seq identity with SEQ ID NO:1 in 518 bp overlap, 97.5 % aa seq identity with SEQ ID NO:2 in 119 aa overlap. the whole document</p> <p>---</p>	1-4,9-11
X	<p>DATABASE STRAND 'Online! GenBank; Locus (Accession) W20138, 3 May 1996 (1996-05-03) HILLIER L ET AL.: "zb40d12.r1 Soares_parathyroid_tumor_NbHPA Homo sapiens cDNA clone IMAGE:306071 5', mRNA sequence" XP002145550 Note: 82.0 % nt seq identity with SEQ ID NO:1 in 713 bp overlap, 56.5 % aa seq identity with SEQ ID NO:2 in 147 aa overlap. the whole document</p> <p>---</p>	1-4,9-11
X	<p>DATABASE STRAND 'Online! GenBank; Locus (Accession) AA229752, 21 August 1997 (1997-08-21) STRAUSBERG R: "nc49d1.r1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1011479, mRNA sequence" XP002145551 Note: 99.6 % nt seq identity with SEQ ID NO:1 in 500 bp overlap. the whole document</p> <p>---</p>	1-4,9-11
A	<p>WO 96 29338 A (UNIV CALIFORNIA ;UNIV PENNSYLVANIA (US); MILLER WALTER L (US); LIN) 26 September 1996 (1996-09-26) Note: 25.6 % aa seq identity with SEQ ID NO:2 in 195 aa overlap. the whole document</p> <p>---</p>	1-23,26
A	<p>US 5 872 230 A (CLARK BARBARA J ET AL) 16 February 1999 (1999-02-16) Note: 26.0 % aa seq identity with SEQ ID NO:2 in 196 aa overlap. the whole document</p> <p>---</p>	1-23,26
	<p>---</p> <p>---/--</p>	

INTERNATIONAL SEARCH REPORT

International Application No

PCT/IL 00/00252

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>SUGAWARA T ET AL.: "Human steroidogenic acute regulatory protein: functional activity in COS-1 cells, tissue-specific expression, and mapping of the structural gene to 8p11.2 and a pseudogene to chromosome 13. "</p> <p>PROC. NAT'L.ACAD. SCI. USA, vol. 92, no. 11, 23 May 1995 (1995-05-23), pages 4778-4782, XP002145548</p> <p>Note: 25.6 % aa seq identity with SEQ ID NO:2 in 195 aa overlap.</p> <p>the whole document</p> <p>---</p>	1-23,26
A	<p>KALLEN C B ET AL.: "Unveiling the mechanism of action and regulation of the steroidogenic acute regulatory protein"</p> <p>MOLECULAR AND CELLULAR ENDOCRINOLOGY, vol. 145, no. 1-2, 25 October 1998 (1998-10-25), pages 39-45, XP000900751</p> <p>cited in the application</p> <p>the whole document</p> <p>---</p>	
P,X	<p>DATABASE STRAND 'Online! GenBank;</p> <p>Locus HSM801955, Accession AL137657, 15 January 2000 (2000-01-15)</p> <p>KOEHRER K ET AL.: "Homo sapiens mRNA; cDNA DKFZp434G2416; partial cds"</p> <p>XP002145552</p> <p>Note: 99.4 % nt seq identity with SEQ ID NO:1 in 994 bp overlap, 98.2 % aa seq identity with SEQ ID NO:2 in 168 aa overlap.</p> <p>page 1, line 39,40</p> <p>---</p>	1-23,26
T	<p>STOCCO D M: "Intramitochondrial cholesterol transfer"</p> <p>BIOCHIMICA ET BIOPHYSICA ACTA, vol. 1486, no. 1, 26 June 2000 (2000-06-26), pages 184-197, XP000937443</p> <p>the whole document</p> <p>-----</p>	

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 24,25

Remark: Present claims 24 and 25 refer to an activator and a deactivator, respectively, of the polypeptide of claim 5, without giving a true technical characterisation. Moreover, no specific compounds are defined in the application. In consequence, the scope of said claims is ambiguous and vague, and their subject-matter is not sufficiently disclosed and supported (Art. 5 and 6 PCT). No search can be carried out for such purely speculative claims whose wording is, in fact, a mere recitation of the results to be achieved.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/IL 00/00252

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9820165	A	14-05-1998	EP 0941366 A	15-09-1999
WO 9629338	A	26-09-1996	US 5807678 A	15-09-1998
			AU 5427596 A	08-10-1996
			JP 2000060581 A	29-02-2000
			JP 10501702 T	17-02-1998
US 5872230	A	16-02-1999	NONE	